A Potential Function for Computer Simulation Studies of Proton Transfer in Acetylacetone

KONRAD HINSEN and BENOÎT ROUX*

Departments of Physics & Chemistry, Université de Montréal, C.P. 6128, succ. Centre-Ville, Montreal, Quebec H3C 3J7, Canada

Received 28 March 1996; accepted 28 May 1996

ABSTRACT

A potential energy model is developed to study the intramolecular proton transfer in the enol form of acetylacetone. It makes use of the empirical valence bond approach developed by Warshel to combine standard molecular mechanics potentials for the reactant and product states to reproduce the interconversion between these two states. Most parameters have been fitted to reproduce the key features of an *ab initio* potential surface obtained from 4-31G* Hartree–Fock calculations. The partial charges have been fitted to reproduce the electrostatic potential surface of 6-31G* Hartree–Fock wave functions, subject to total charge and symmetry constraints, using a fitting procedure based on generalized inverses. The resulting potential energy function reproduces the features most important for proton transfer simulations, while being several orders of magnitude faster in evaluation time than *ab initio* energy calculations.

Introduction

Proton transfer reactions are important in several areas of chemistry and biology and have therefore been studied intensively.[1,2] Intramolecular proton transfer reactions are particularly attractive from a theoretical point of view, because they provide a well-defined situation without the complications associated with the translational diffusion of multiple donors and acceptors in a solvent. However, most theoretical studies have been based on simplified model systems[3–5] that do not allow quantitative comparisons with experimental data. Nevertheless, various numerical techniques allowing a rigorous treatment of quantum effects in realistic systems have been developed in recent years. Quantum statistical calculations based on discretized Feynman path integrals[6,7] can be applied to systems with a considerable number of degrees of freedom, and dynamical methods such as wavepacket propagation[8] can be used on a quantum subsystem coupled to a classical environ-

* Author to whom all correspondence should be addressed.
E-mail: rouxb@ere.umontreal.ca

CCC 0192-8651 / 97 / 030368-13

ment that is treated with classical molecular dynamics techniques. To apply such methods to a realistic model for a real proton transfer system, it is necessary to have a sufficiently accurate potential energy surface that can be evaluated in an acceptable amount of CPU time. It is the aim of this work to provide such a potential energy surface model for the study of proton transfer in acetylacetone.

The most rigorous potential energy surfaces available today are those obtained from *ab initio* calculations based on various levels of Hartree–Fock MP2, MP3, or density functional theory approaches.[9,10] Although such energy surfaces have been used to perform classical molecular dynamics simulations for small systems,[11,12] the evaluation of the energy and the forces for a given nuclear configuration is much too expensive to be used at each time step of a fully or partly quantum-mechanical simulation, even in the absence of solvent. An important feature of quantum simulation methods, such as path integration or wavepacket propagation, is that they require the frequency evaluation of the potential energy at many more nuclear configurations than classical simulations. Semiempirical quantum chemistry methods are somewhat faster and have been used in classical simulations,[13] but are still impractical for a quantum treatment of the nuclei. This leaves the construction of an appropriate potential energy model function with parameters obtained by some fitting procedure as the only viable option. It should be pointed out that this approach is not only several orders of magnitude faster than *ab initio* calculations, but might also turn out to be more accurate than a low-level *ab initio* calculation, provided that high-level calculations have been used in parameter fitting.

The molecule we have chosen for our studies of intramolecular proton transfer is the enol form of acetylacetone, whose structure is given in Figure 1. The proton H_a can be bound to either of the two oxygen atoms and the molecule can undergo interconversion between these two states. The fact that the reactant and product states are mirror images of each other simplifies the construction of a model potential significantly. A more popular molecule in studies of intramolecular proton transfer has been malonaldehyde, which differs from acetylacetone by the absence of the two methyl groups. However, malonaldehyde is not stable in common solvents, whereas acetylacetone is stable in a wide range of polar and nonpolar solvents. Since we intend to study solvated systems in the future, we have chosen acetylacetone for our studies.

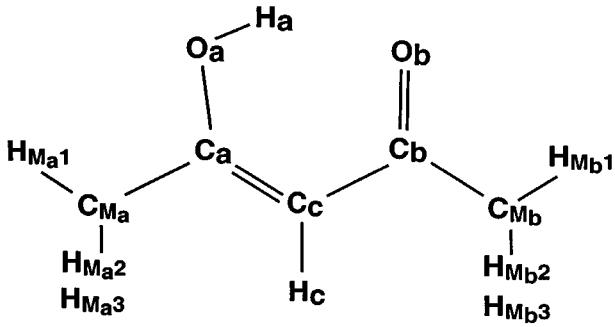

FIGURE 1. The minimum energy conformation of acetylacetone.

Our potential energy model consists mostly of standard bond, angle, and dihedral potentials as used in molecular mechanics applications. The ability to form and break bonds is introduced by combining the reactant and product potentials using the empirical valence bond (EVB) approach developed by Warshel and coworkers.[14–17] We also used a Lennard–Jones type potential for the O—H bond to improve the accuracy of the potential energy surface for the transferring proton (all other bond potentials are harmonic). All parameters of the resulting model are fitted to the results of *ab initio* energy calculations. In this way, we obtain a potential energy model that reflects the salient features of the *ab initio* potential while being hardly more expensive to evaluate than standard molecular mechanics potential terms.

Construction of a Potential Function

FUNCTIONAL FORM

The essential problem with the standard potential functions used in classical simulations is that they do not permit the breakup and formation of chemical bonds. To obtain a meaningful description of the potential along the reaction path, we used the empirical valence bond (EVB) approach,[14] which describes the molecule as a coupled two-state system. The two states correspond to the reactant and the product state of the proton transfer reaction; in the case of acetylacetone, they are mirror images of each other. For one configuration $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2 \cdots)$ of the molecule, the EVB potential

is defined as the lower eigenvalue of the ''Hamiltonian'':

$$H = \begin{pmatrix} U_1(\mathbf{R}) & \epsilon \\ \epsilon & U_2(\mathbf{R}) \end{pmatrix} \quad (1)$$

which is given by

$$E_{\mathrm{EVB}}(U_1, U_2, \epsilon) = \frac{U_1 + U_2}{2} - \sqrt{\left(\frac{U_1 + U_2}{2}\right)^2 - U_1 U_2 + \epsilon^2} \quad (2)$$

This leads to a double-well potential where the wells correspond to the reactant and product states (although somewhat modified) and the height of the barrier between them is controlled by the coupling parameter, ϵ, whose value must be found as part of a fitting procedure. The reactant and product potentials, U_1 and U_2, consist of standard molecular mechanical bond, angle, dihedral, and nonbonded potentials; we use the functional form of the CHARMM potential,[18] in which the various contributions are:

$$E_{\mathrm{bond}}(r) = k_r(r - r_{\mathrm{min}})^2 \quad (3)$$

$$E_{\mathrm{angle}}(\theta) = k_\theta(\theta - \theta_{\mathrm{min}})^2 \quad (4)$$

$$E_{\mathrm{dihedral}}(\phi) = |k_\phi| - k_\phi \cos(n_\phi \phi - \phi_{\mathrm{min}}) \quad (5)$$

$$E_{\mathrm{improper}}(\omega) = k_\omega(\omega - \omega_{\mathrm{min}})^2 \quad (6)$$

Some angles have an additional Urey–Bradley potential of the form:

$$E_{\mathrm{UB}}(r) = k_{\mathrm{UB}}(r - r_{\mathrm{min}})^2 \quad (7)$$

where r is the distance between the two outer atoms.

For the H_a—O_a and H_a—O_b bonds, the anharmonic character of the potential surface was significant and it was found necessary to have a more accurate representation than a simple harmonic potential. We used a Lennard–Jones-type potential parameterized as:

$$E_{\mathrm{LJ}}(r) = 2E_0\left[\frac{1}{2}\left(\frac{r_0}{r + r_0 - r_{\mathrm{min}}}\right)^{12} - \left(\frac{r_0}{r + r_0 - r_{\mathrm{min}}}\right)^{6}\right] \quad (8)$$

with

$$r_0 = \sqrt{\frac{36E_0}{k_r}}$$

Here r_{min} is the equilibrium bond length, E_0 is the energy at this bond length (the energy is zero at $r = \infty$), and k_r is the force constant of an harmonic approximation to the Lennard–Jones potential around $r = r_{\mathrm{min}}$.

The nonbonded potential is defined by standard partial charges and Lennard–Jones interactions:

$$E_{\mathrm{Coulomb}}(r_{ij}) = \frac{q_i q_j e^2}{r_{ij}} \quad (9)$$

$$E_{\mathrm{LJ}}(r_{ij}) = 2E_0\left[\frac{1}{2}\left(\frac{R_{\mathrm{min}}^{(ij)}}{r_{ij}}\right)^{12} - \left(\frac{R_{\mathrm{min}}^{(ij)}}{r_{ij}}\right)^{6}\right] \quad (10)$$

As usual, the nonbonded potential is not calculated for nearest and next-to-nearest neighbors along a chain of bonds (1–2 and 1–3 interactions).

PARAMETERIZATION AND OPTIMIZATION PROCEDURE

The parameters for the model described in the previous section were adjusted to reproduce as well as possible an *ab initio* potential energy surface. The following criteria were used in the fitting procedure:

- ground state geometry;
- electrostatic potential surface around the molecule;
- vibrational frequencies;
- energy difference between ground and transition state;
- potential energy surface for the transferring proton with all other atoms fixed in the ground and transition states; and
- energy as a function of distorted ground state geometry.

The reference ground state geometry was obtained from an energy minimization on a Hartree–Fock potential energy surface using Gaussian 90.[19] The transition state geometry was obtained similarly by minimizing the potential energy under constraints imposing C_{2V} symmetry.

The choice of basis sets in Hartree–Fock calculations is always a compromise between accuracy and computational effort. We have chosen the 4-31G* basis set for everything except the evalua-

tion of the electrostatic potential surface; this basis set has been used successfully in other studies.[20] The electrostatic potential energy surface used for fitting partial charges was obtained with a 6-31G* basis set, whose suitability for such applications has been established by Carlson et al.[21]

Nonbonded Interactions

The first step in the fitting procedure was the determination of the parameters for nonbonded interactions. The Lennard–Jones parameters, which describe mainly intermolecular interactions, were taken from parameters for similar molecules in the CHARMM PARAM22 force field.[22] The partial charges were determined by fitting to the electrostatic potential surface of the molecule. The reference data were obtained from 6-31G* Hartree–Fock calculations with Gaussian 90,[19] using the ground and transition state geometries from the 4-31G* minimizations.

The fitting procedure that was used differs significantly from others published previously.[23–28] The general aim is to find the set of partial charges which best reproduces the *ab initio* electrostatic potential at a set of points which are thought to be representative of the accessible region during an interaction with another molecule. In early approaches, a set of points evenly distributed on a grid located in the accessible volume around the molecule has been used.[23, 24, 26, 27] Using a grid implies an arbitrary choice of directions for its axes, and sometimes the results are strongly dependent on that choice. A better procedure is to select points on several surfaces that have a fixed distance from the van der Waals molecular surface defined by the atomic radii.[25, 28] In the present work, we have used a set of points randomly distributed in the accessible volume. This procedure is much simpler, because it does not require the construction of the molecular surface, and has the additional advantage of making it easy to test the convergence of the fitted charges by adding more points. Furthermore, a scheme based on generalized matrix inverses was used both for the solution of the least-squares problem and for handling constraints. This scheme is conceptually clearer than the more common approach based on Lagrange multipliers and solution of the normal equations, and it also helps in analyzing problematic situations where the data points are not sufficient to determine a set of point charges unambiguously. This is described in detail in the Appendix.

First, we select a set of points, $\mathbf{r}_j$ ($j = 1 \cdots n$), at which the potential, $\phi_{\text{ref}}(\mathbf{r}_j)$, is to be evaluated. The points are randomly chosen in a region around the molecule in which each point is at least 3 Å away from the heavy atoms (C and O), but no more than 5 Å from any atom. This covers the solvent-accessible region of the molecule. Within this region, the points have a uniform distribution. It was found that 1000 points are sufficient to obtain convergence to an accuracy of three digits for acetylacetone.

Denoting the positions of the nuclei by $\mathbf{R}_i$ ($i = 1 \cdots N$), and the partial charges to be determined by q_i, we can write the potential of the partial charges at the evaluation points as:

$$\phi_q(\mathbf{r}_j) = \sum_i A_{ji} q_i \tag{11}$$

with:

$$A_{ji} = \frac{1}{|\mathbf{r}_j - \mathbf{R}_i|} \tag{12}$$

It should be noted that the rows of **A** corresponding to different j need not refer to the same nuclear configuration; it is possible to obtain optimal fits for several conformations by combining them in one fitting problem.

The aim of the least-squares procedure is to find the values of the charges, q_i, that minimize the quantity:

$$\chi^2 = \sum_j \left[\phi_q(\mathbf{r}_j) - \phi_{\text{ref}}(\mathbf{r}_j) \right]^2 \tag{13}$$

However, the set of charges must also obey the imposed constraints, such as fixed total charge, fixed total dipole, or equality of certain charges due to symmetry. All constraints of interest are linear in the charges and can therefore be written in the form:

$$\sum_i B_{ki} q_i = c_k, \qquad k = 1 \cdots n_c \tag{14}$$

where n_c is the number of constraints. For example, if constraint k is to define the total charge to be zero, all B_{ki} would be one and c_k would be zero. A method for finding the minimum of χ^2 subject to constraints of the form given in eq. (14) based on generalized matrix inverses is described in the Appendix.

We have applied this procedure with the following constraints on the charges:

- The total charge of the molecule must be zero.
- The charges of all hydrogens in the two methyl groups are identical.
- The charges in the reactant and product ground states are the same, i.e., $O_a = O_b$, $C_{M_a} = C_{M_b}$, and $C_a = C_b$.

The last condition is an approximation which makes it possible to calculate the electrostatic interactions of the acetylacetone molecules with other molecules only once outside of the EVB relation eq. (2), because $E_{\mathrm{EVB}}(U_1 + U_{\mathrm{el}}, U_2 + U_{\mathrm{el}}) = U_{\mathrm{el}} + E_{\mathrm{EVB}}(U_1, U_2)$. To check the effect of imposing the constraints, alternative charge values were calculated with only the total charge constraint. The differences between the two sets of charges were found to be small; the average absolute difference is $0.026e$, the highest is $0.1e$ (for C_b). To see how well these charges represent the transition state, a separate fit was performed using the transition state geometry. Again the differences to the ground state charges were found to be small with an average difference of $0.1e$ and a maximum of $0.3e$ (for C_c).

Internal Energy Terms and Coupling Parameter

The next step in the fitting procedure involved all parameters that affect the ground state geometry, i.e., the bond and bond angle parameters as well as the coupling parameter, ϵ. To simplify this process, the $H_M—C_M—H_M$ angle was excluded. This angle is the only one that occurs in a nonplanar substructure; adjusting its parameters would complicate the fitting procedure without leading to a significant improvement, because there is no reason to assume that details of the methyl group potential will affect proton transfer. The force constant for this angle was taken from the CHARMM PARAM22 force field,[22] and its equilibrium angle was taken from the ground state reference geometry.

The force constants for the remaining bond and bond angle energy terms were obtained from the force constant matrix resulting from the *ab initio* calculations for the ground state. The equilibrium bond length and angle parameters were then adjusted together with the EVB coupling parameter, ϵ, and all the parameters of the Lennard–Jones potential for the $H_a—O_a$ bond [see eq. (8)] to obtain good agreement with the *ab initio* results for the ground state geometry, the energy difference between ground and transition state, and the potential energy surface for H_a with all other atoms in their ground state positions. This was the most time-consuming step of the fitting procedure. The bond lengths and angles from the ground state *ab initio* calculations were used as initial values, together with a value for ϵ that gave the correct energy difference to the transition state. Reasonable guesses were used for the $H_a—O_a$ parameters. All parameters were then improved in an iterative manner, adjusting the geometric parameters and ϵ in turns. Initially, least-squares fits using a linearization of the potential based on finite-difference approximations were used, but in the later stages the finite-difference approximation became insufficient and the final adjustments had to be made manually.

Because both the ground state and the transition state are planar, the force constants for the dihedral angle energy terms have no influence on the two geometries and therefore on the quantities used so far in the fitting procedure. They can be obtained independently from *ab initio* evaluations of the energy of torsional deformation. The energy of several distorted configurations (obtained from the ground state geometry by changing individual dihedral angles) was obtained in 4-31G* approximation: torsion around the $C_a—C_c$ bond, torsion around the $C_c—C_b$ bond, and torsion of only the methyl groups around these axes. These were used to fit the dihedral energy terms involving the axes $C_a—C_c$ and $C_c—C_b$ and the improper dihedrals around C_a and C_b.

The force constants for the dihedral terms related to the out-of-plane motion of the proton H_a were again taken from the CHARMM PARAM22 force field.[22] The remaining dihedral angle potential, the one for the dihedral angles describing the hydrogen atoms in the methyl groups, was adjusted to reproduce the lowest vibrational frequencies, corresponding to rotational vibrations of the methyl groups, to the extent possible. The reference frequencies were again obtained from 4-31G* Hartree–Fock calculations.

Results and Discussion

The reference geometry of the reactant/product ground state, obtained from an *ab initio* energy

minimization, is given in Table I, together with the minimum-energy geometry of the fitted potential function. The excellent agreement is not surprising, given that the ground state geometry was fitted with a sufficiently large set of free parameters. It is more interesting to compare the geometries of the transition state, given in Table II. The transition state geometry was not used at all in the fitting procedure, and yet it shows a good agreement. The bond lengths between the heavy atoms are reproduced well, and also the angles involving the proton H_a. In addition, the changes in the internal coordinates as the molecule passes from the reactant to the transition state have the correct direction; i.e., the bonds that get shorter according to the reference potential energy surface also get shorter according to the fitted potential energy surface. The most important deviation from the reference geometry is in the $O_a—O_b$ distance, which is too large by 0.1 Å. This is a consequence of accumulated small deviations in the angles along the ring. The energy difference between ground state and transition state is 9.44 kcal/mol from both the *ab initio* and the fitted potential energy surface.

TABLE I.
Ground State Geometry for *Ab Initio* and Fitted Potential.

	Ab initio	Model potential
Distance [Å]		
$H_a—O_a$	0.9645	0.9645
$H_a—O_b$	1.7922	1.7930
$O_b—O_a$	2.6243	2.6251
$O_a—C_a$	1.3139	1.3139
$C_a—C_c$	1.3462	1.3462
$C_c—C_b$	1.4533	1.4535
$C_b—O_b$	1.2115	1.2115
$C_a—C_{M_a}$	1.5020	1.5020
$C_{M_a}—H_{M_a1}$	1.0800	1.0824
$C_{M_a}—H_{M_a2}$	1.0833	1.0806
$C_{M_a}—H_{M_a3}$	1.0833	1.0806
$C_c—H_c$	1.0726	1.0725
$C_b—C_{M_b}$	1.5094	1.5094
$C_{M_b}—H_{M_b1}$	1.0798	1.0819
$C_{M_b}—H_{M_b2}$	1.0852	1.0802
$C_{M_b}—H_{M_b3}$	1.0852	1.0802
Angle [degrees]		
$H_a—O_a—C_a$	108.852	108.850
$O_a—C_a—C_c$	123.738	123.742
$C_a—C_c—C_b$	121.782	121.795
$C_c—C_b—O_b$	122.512	122.517
$O_a—C_a—C_{M_a}$	113.207	113.218
$C_c—C_a—C_{M_a}$	123.056	123.040
$C_a—C_{M_a}—H_{M_a1}$	109.900	110.444
$C_a—C_{M_a}—H_{M_a2}$	110.402	109.148
$C_a—C_{M_a}—H_{M_a3}$	110.402	109.148
$C_a—C_c—H_c$	119.025	119.017
$C_b—C_c—H_c$	119.193	119.188
$C_c—C_b—C_{M_b}$	117.269	117.269
$O_b—C_b—C_{M_b}$	120.219	120.214
$C_b—C_{M_b}—H_{M_b1}$	109.690	109.384
$C_b—C_{M_b}—H_{M_b2}$	110.160	109.170
$C_b—C_{M_b}—H_{M_b3}$	110.160	109.170

An important feature of a potential energy surface for quantum-mechanical proton transfer simulations is the shape of the potential energy surface for the proton with the remaining nuclei in fixed positions. Due to the significant delocalization of the proton (the de Broglie thermal wave length of a proton is around 0.5 Å at 300 K) this whole surface is expected to have an important influence on the behavior of the system. Reproducing the main features of the whole surface is obviously a much more difficult task than obtaining a good ground state geometry. As shown in Figure 2, the proton energy surface is at least qualitatively correct in the fixed reactant configuration. The position and depth of the second higher-energy well are somewhat different, but the energy difference between the wells is so large that the second well is not expected to play an important role. The proton energy surface in the fixed transition state configuration of the fitted potential is shown in Figure 3. Again, the main features of the potential energy surface are well reproduced, although the fitted potential lacks the shallow double-well character observed in the *ab initio* potential. This detail of the *ab initio* surface could not be reproduced exactly with the current potential function. Nevertheless, because the barrier in the *ab initio* surface is low (1.7 kcal/mol), we do not expect this to be a serious deficiency. A numerical solution of the one-dimensional Schrödinger equation for this double-well potential indicates that the ground state wave function has a barely noticeable two-peak structure. Moreover, there is an asymmetric double well for both the *ab initio* and the fitted potential for all other configurations (this is also the case for any small deviations in the neighborhood of the transition symmetric state). Because the occurrence of the perfectly symmetric transition state (a rather singular configuration) is an extremely rare event, such small detail of the potential energy surface should not have a strong influence on complex dynamical quantities involving larger regions of configuration space.

Figure 4 shows the energy as a function of various torsional deformations from the ground

TABLE II.
Transition State Geometry for Ab Initio and Fitted Potential.

	Ab initio	Model potential
Distance [Å]		
$H_a - O_a$, $H_a - O_b$	1.1919	1.2444
$O_a - O_b$	2.3307	2.4317
$O_a - C_a$, $O_b - C_b$	1.2594	1.2536
$C_a - C_{M_a}$, $C_b - C_{M_b}$	1.5017	1.5045
$C_a - C_c$, $C_c - C_b$	1.3954	1.3886
$C_{M_a} - H_{M_a1}$, $C_{Mb} - H_{M_b1}$	1.0802	1.0821
$C_{M_a} - H_{M_a2}$, $C_{M_b} - H_{M_b2}$	1.0838	1.0804
$C_{M_a} - H_{M_a3}$, $C_{M_b} - H_{M_b3}$	1.0838	1.0804
$C_c - H_c$	1.0707	1.0726
Angle [degrees]		
$O_a - H_a - O_b$	155.741	155.400
$H_a - O_a - C_a$, $H_a - O_b - C_b$	103.185	101.763
$O_a - C_a - C_c$, $C_c - C_b - O_b$	120.540	120.408
$C_a - C_c - C_b$	116.809	120.256
$O_a - C_a - C_{M_a}$, $O_b - C_b - C_{M_b}$	117.738	117.872
$C_c - C_a - C_{M_a}$, $C_c - C_b - C_{M_b}$	121.722	121.720
$C_a - C_{M_a} - H_{M_a1}$, $C_b - C_{M_b} - H_{M_b1}$	109.432	109.729
$C_a - C_{M_a} - H_{M_a2}$, $C_b - C_{M_b} - H_{M_b2}$	110.283	109.189
$C_a - C_{M_a} - H_{M_a3}$, $C_b - C_{M_b} - H_{M_b3}$	110.284	109.189
$C_a - C_c - H_c$, $C_b - C_c - H_c$	121.596	119.872

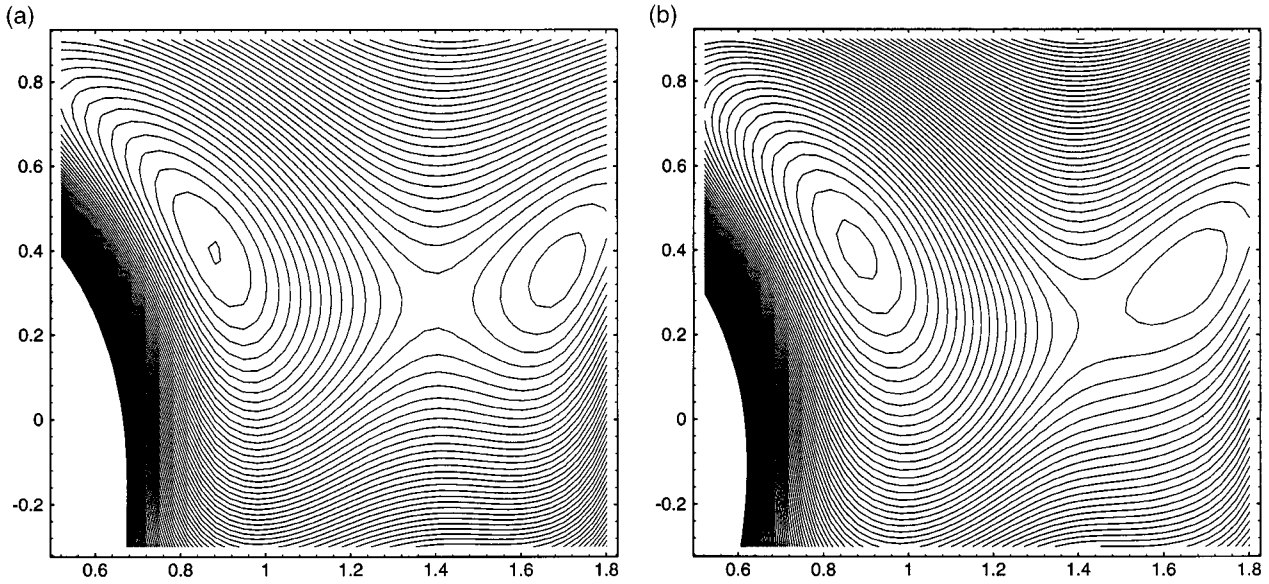

FIGURE 2. The potential energy surface for H_a in the plane of the heavy atoms, which are in the ground state configuration. The axes are labeled with the distances (in angstroms) from O_a along the principal axes of inertia (same orientation as for Fig. 1). The energy difference between two contour lines is 2 kcal / mol. (a) *Ab initio* potential (the contour plot is generated from 16 by 16 energy values. (b) Fitted model potential (the contour plot is generated from 61 by 71 energy values).

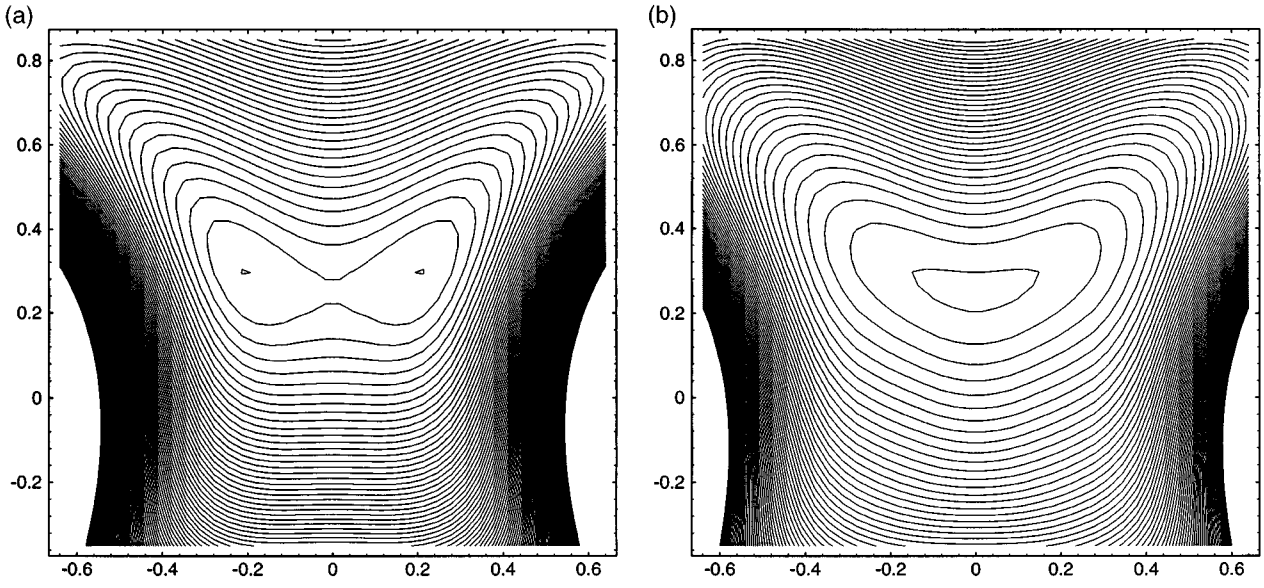

FIGURE 3. The potential energy surface for H_a in the plane of the heavy atoms, which are in the transition state configuration. The axes are labeled with the distances (in angstroms) from the point between O_a and O_b along and perpendicular to the O_a — O_b axis. The energy difference between two contours is 2 kcal / mol. (a) *Ab initio* potential (the contour plot is generated from 16 by 17 energy values). (b) Fitted model potential (the contour plot is generated from 61 by 71 energy values).

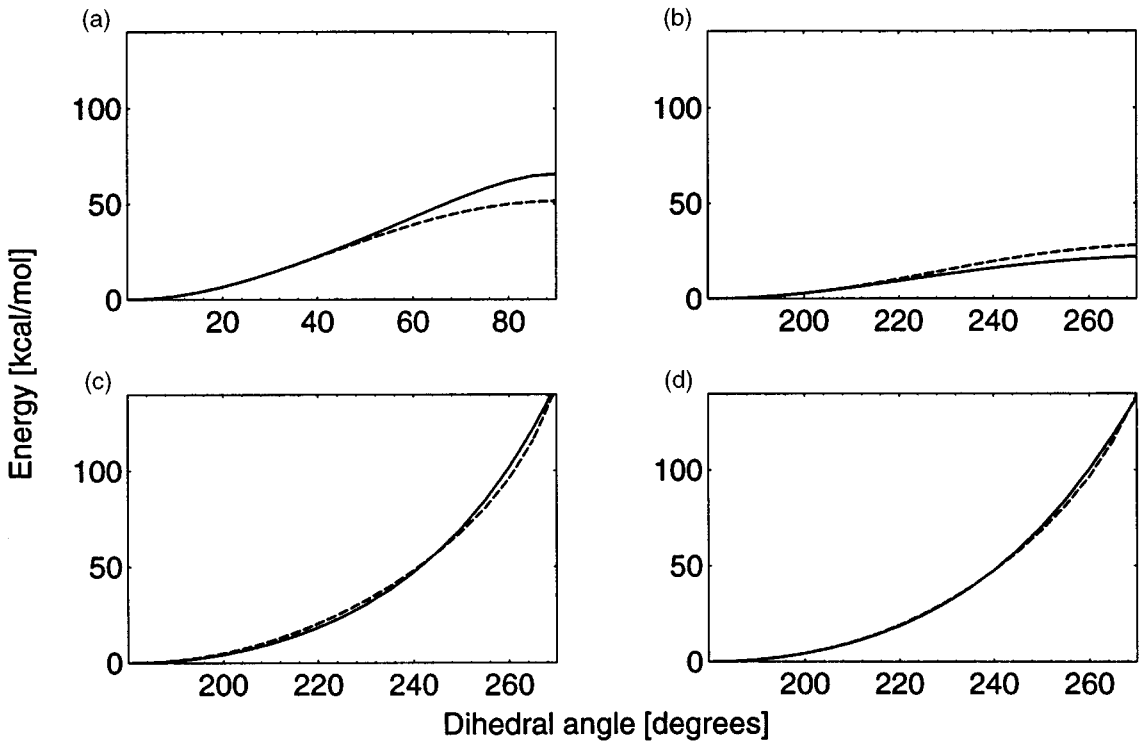

FIGURE 4. The energy as a function of torsional deformation from the ground state. The solid lines represent the *ab initio* potential, and the dashed lines the fitted model potential. (a) Rotation of H_a, O_a, C_a, and the methyl group around C_{M_a} around the axis C_a — C_c. (b) Rotation of O_b, C_b, and the methyl group around C_{M_b} around the axis C_c — C_b. (c) Rotation of the methyl group around C_{M_a} around the axis C_a — C_c. (d) Rotation of the methyl group around C_{M_b} around the axis C_c — C_b.

TABLE III.
Ten Lowest Vibrational Frequencies per Centimeter.

Ab initio	Model potential
93.0103	88.62
122.2650	124.68
160.3442	136.21
199.3143	169.58
239.9272	287.55
390.0118	321.70
417.6479	449.05
534.9204	471.93
602.2793	499.37
696.7507	595.64

state. It is evident that the fit is very good for the range of angles that can be expected to occur under the influence of thermal fluctuations. Finally, the first ten vibrational frequencies are compared in Table III. Because only one parameter was left to adjust the low-frequency vibrations, one would not expect a very good agreement, but differences are nevertheless not very large.

The potential parameters are listed in Tables IV–VII. Table IV gives the Lennard–Jones parameters, which are taken from the CHARMM PARAM22 force field,[22] and the partial charges obtained from fitting to the *ab initio* electrostatic potential for the ground state geometry. The parameters for the bonded interactions are given in Tables V (bond), VI (bond angle), and VII (dihedral angle).

The fitted value for the coupling parameter ϵ is 101.33 kcal/mol. The parameters for the $H_a—O_a$ bond [see eq. (8)] are $E_0 = 500$ kcal/mol, $r_{\min} =$ 0.89135 Å, and $k_r = 1000$ kcal/mol Å^2; these values produce the best possible agreement of the proton potential energy surface (Figures 2b and 3b).

TABLE IV.
Lennard–Jones Parameters and Partial Charges [see Eqs. (9) and (10)].[a]

Atom	q [e]	E_0 [kcal/mol]	$R_{\min}/2$ [Å]
C_a	0.825	−0.07	1.9924
C_c	−0.930	−0.07	1.9924
C_b	0.825	−0.07	1.9924
O_a	−0.615	−0.1521	1.77
O_b	−0.615	−0.1521	1.77
H_a	0.443	−0.046	0.2245
H_c	0.245	−0.03	1.3582
C_M	−0.602	−0.08	2.06
H_M	0.171	−0.022	1.32

[a] The Lennard–Jones parameters are taken from the CHARMM PARAM22 force field[22] for equivalent atoms.

TABLE V.
Force Constants for Bonds [see Eq. (3)].

Bond	$r_{\min}$ [Å]	k_r [kcal/mol Å^2]
$O_a—C_a$	1.3413	612.0
$C_a—C_c$	1.3639	730.1
$C_c—C_b$	1.4805	429.3
$C_b—O_b$	1.2146	1015.7
$C_{M_a}—C_a$	1.5363	386.5
$C_b—C_{M_b}$	1.5466	358.6
$C_c—H_c$	1.08165	450.6
$C_M—H_M$	1.0850	424.0

It should be noted that our use of the EVB method differs from previous applications.[14,15] In these applications, the energy surfaces for the ''pure'' states are given as inputs; e.g., they were obtained from semiempirical or *ab initio* calculations, and the EVB method then provides a kind of interpolation to produce an energy surface that allows reactions. The off-diagonal elements in eq. (1) are either also known or obtained via some approximation as a function of the diagonal elements. Therefore, they are, in general, configuration-dependent. In the present case, the EVB method is used to obtain a model that allows bond breaking and formation from a model that does not. No information about the off-diagonal coupling is available. Because the model already

TABLE VI.
Force Constants for Angles [see Eq. (4)].

Angle	$\theta_{\min}$ [degrees]	k_θ [kcal/mol]
$H_a—O_a—C_a$	115.93	88.3
$O_a—C_a—C_c$	122.95	142.9
$C_a—C_c—C_b$	119.62	177.2
$C_c—C_b—O_b$	120.59	140.0
$C_{M_a}—C_a—O_a$	112.29	79.0
$C_{M_a}—C_a—C_c$	124.76	75.6
$C_{M_b}—C_b—O_b$	119.48	98.2
$C_{M_b}—C_b—C_c$	119.93	59.8
$H_c—C_c—C_a$	120.04	41.8
$H_c—C_c—C_b$	120.34	48.3
$H_M—C_M—H_M$	108.80	35.5
$H_M—C_{M_a}—C_a$	111.00	49.7
$H_M—C_{M_b}—C_b$	111.0	46.0

TABLE VII.
Force Constants for Dihedral Angles and Improper Dihedral Angles [see Eqs. (5) and (6)].

Dihedral	ϕ_{min}, ω_{min} [degrees]	k_ϕ, k_ω [kcal / mol]	n_ϕ
$H_a—O_a—C_a—C_c$	180	3.0	2
$O_a—C_a—C_c—C_b$	180	10.0	2
$C_a—C_c—C_b—O_b$	180	1.0	2
$C_M—C_a—C_c—C_b$	180	2.0	2
$C_a—C_c—C_b—C_M$	180	5.5	2
$H_a—O_a—C_a—C_M$	180	1.0	2
$O_a—C_a—C_c—H_c$	180	5.5	2
$O_b—C_b—C_c—H_c$	180	1.5	2
$C_M—C_a—C_c—H_c$	180	5.5	2
$C_M—C_b—C_c—H_c$	180	1.5	2
$H_M—C_M—C_a—O$	180	0.1	3
$H_M—C_M—C_b—O$	180	0.1	3
$C_a—H_c—C_b—C_c$	0	10.0	0
$C_b—H_c—C_a—C_c$	0	10.0	0
$O_a—C_{M_a}—C_c—C_a$	0	90.0	0
$O_b—C_{M_b}—C_c—C_b$	0	90.0	0

contains many fit parameters, the simplest working approximation for the off-diagonal element was chosen. As our results show, a good fit can be obtained with this choice, and therefore, more complicated options were not investigated. Fitted parameters in an EVB potential have been used before in the context of free energy barrier calculations.[29]

As one would naively expect, the structural properties of the minimum-energy configuration obtained with the full EVB potential energy surface resembles the one found with the ''pure'' state potential, U_1; the bond lengths differ by about 3% and the angles by up to 5%. However, despite such small structural differences, the product and reactant EVB states differ significantly from the ''pure'' states described by the surfaces U_1 and U_2 [see eq. (2)]. This is an immediate consequence of the strong coupling between the two pure states due to the off-diagonal element, ϵ. It follows that the properties of the EVB potential, such as the vibrational frequencies, cannot be readily deduced from those of the pure states. Of particular importance, the potential energy surface seen by the transferring proton (which is most directly dependent on the $O_a—O_b$ distance) is substantially affected by the coupling between the two pure states. The energy difference, ΔE, between ground and transition state depends strongly on the value of ϵ; its derivative is given by $\partial \Delta E / \partial \epsilon = -0.74$. To obtain a meaningful potential energy surface it is therefore essential to parameterize the model, including the value of ϵ, very carefully.

The choice of the EVB approach to model bond formation leads to the symmetry restriction on the charges for practical reasons, as discussed in the ''Nonbonded Interactions'' subsection. However, this restriction is equivalent to the condition that the partial charges in the final potential energy surface be independent of the position of the nuclei, which is a condition that is almost always imposed on partial charges in empirical force fields. Moreover, we have shown that its effect on the resulting charges is small.

Summary

We have developed a potential function for an atomic model of acetylacetone that is detailed enough to allow realistic proton transfer studies, yet simple enough to be evaluated rapidly, in contrast to *ab initio* potentials. The model is based on the EVB method[14] and standard molecular mechanics harmonic potentials for bonds, angles, and dihedrals, plus a Lennard–Jones potential for the $H_a—O_a$ bond. The parameters were adjusted to match as closely as possible the main features (geometry, energy barrier, energy surface for the proton) of a 4-31G* *ab initio* potential and the electrostatic potential of a 6-31G* Hartree–Fock wave function.

We have used our potential function in a discretized Feynman path integral calculation of the potential of mean force and rate constant for the proton transfer in the gas phase and in various solvents; the results will be published elsewhere. Another promising area of application that we are working on is the use of wavepacket propagation techniques to study the real-time dynamics of proton transfer. We hope that our work will encourage others to apply advanced numerical techniques to our realistic model potentials, and to develop similar potential functions for other systems.

Appendix: Charge Fitting with Generalized Inverses

In this appendix we show how the charges that minimize eq. (13), subject to the constraints given in eq. (14), can be obtained in the form of an explicit expression. This expression can be evalu-

ated accurately and noniteratively, and the evaluation procedure can help to locate problematic situations where the charges cannot be determined unambiguously, and even give hints about how to improve the fit by adding more data points.

For brevity, we write eq. (11) in the form:

$$\mathbf{Aq} = \mathbf{p}_q \tag{15}$$

[the ith element of $\mathbf{p}_q$ contains $\phi_q(\mathbf{r}_i)$] and eq. (14) as:

$$\mathbf{Bq} = \mathbf{c} \tag{16}$$

We denote the vector containing the values of the reference potential by $\mathbf{p}_{\mathrm{ref}}$.

Equation (14) is an underdetermined system of linear equations for the partial charges $\mathbf{q}$. It has a solution if and only if:

$$\mathbf{BB}^{+}\mathbf{c} = \mathbf{c} \tag{17}$$

where $\mathbf{B}^{+}$ is the generalized inverse of the matrix $\mathbf{B}$.[30] If this condition is fulfilled, the general solution is given by:

$$\mathbf{q} = \mathbf{B}^{+}\mathbf{c} + (\mathbf{1} - \mathbf{B}^{+}\mathbf{B})\mathbf{q}' \tag{18}$$

for an arbitrary vector, $\mathbf{q}'$; otherwise, the constraints are incompatible and there is no solution. By inserting eq. (18) into eq. (13), we obtain a new least-squares problem for $\mathbf{q}'$:

$$|\mathbf{AB}^{+}\mathbf{c} + \mathbf{APq}' - \mathbf{p}_{\mathrm{ref}}|^2 = min \tag{19}$$

with

$$\mathbf{P} = \mathbf{1} - \mathbf{B}^{+}\mathbf{B} \tag{20}$$

This is equivalent to finding the ''best'' solution (in the least-squares sense) of the system of linear equations:

$$\mathbf{A}'\mathbf{q}' = \mathbf{p}' \tag{21}$$

with:

$$\mathbf{A}' = \mathbf{AP}, \qquad \mathbf{p}' = \mathbf{p}_{\mathrm{ref}} - \mathbf{AB}^{+}\mathbf{c} \tag{22}$$

This system of equations is at the same time overdetermined (there are more equations than unknowns) and underdetermined (the column rank of the matrix $\mathbf{A}'$ is smaller than the number of unknowns, since the matrix $\mathbf{P}$ is in general singular); i.e., there is no exact solution and the least-squares solution is not unique.

The general least-squares solution of eq. (21) is given by:[30]

$$\mathbf{q}' = \mathbf{A}'^{+}\mathbf{p}' + (\mathbf{1} - \mathbf{A}'^{+}\mathbf{A}')\mathbf{x} \tag{23}$$

for an arbitrary vector $\mathbf{x}$; the error of the solution is the same for all $\mathbf{x}$. We obtain the solution with minimal length for $\mathbf{x} = 0$; the reason for choosing the minimal-length solution will be explained in what follows. Having found $\mathbf{q}'$, we insert it into eq. (18) to get the original charges $\mathbf{q}$. It should be noted that the general inverses must be calculated with a numerically stable method, such as singular value decomposition (SVD).

The majority of previously published charge fitting methods[24–28] differ from our scheme on two points: (1) they use Lagrange multipliers to take constraints into account instead of eliminating them as we do in eq. (18); and (2) they obtain the least-squares fit for the charges by solving the normal equations. The second difference is the more important one, and to illustrate the advantage of our approach we will for now assume that no constraints are present; i.e., that the fitting problem is given by finding the least-squares solution of the linear equations $\mathbf{Aq} = \mathbf{p}_{\mathrm{ref}}$.

The method described above gives the solution:

$$\mathbf{q} = \mathbf{A}^{+}\mathbf{p}_{\mathrm{ref}} + (\mathbf{1} - \mathbf{A}^{+}\mathbf{A})\mathbf{x} \tag{24}$$

for arbitrary $\mathbf{x}$. The normal equations for this problem are:

$$\mathbf{A}^{T}\mathbf{Aq} = \mathbf{A}^{T}\mathbf{p}_{\mathrm{ref}} \tag{25}$$

and they can easily be derived by setting the derivative of the least-squares condition with respect to $\mathbf{q}$ equal to zero. Their solution is:

$$\mathbf{q} = (\mathbf{A}^{T}\mathbf{A})^{-1}\mathbf{A}^{T}\mathbf{p}_{\mathrm{ref}} \tag{26}$$

The relation between the two approaches can be seen from a well-known formula for the generalized inverse:

$$\mathbf{A}^{+} = (\mathbf{A}^{T}\mathbf{A})^{-1}\mathbf{A}^{T} \tag{27}$$

which, however, is valid only if the matrix $\mathbf{A}$ has full column rank. Under that condition, $(\mathbf{1} - \mathbf{A}^{+}\mathbf{A})$ is zero, such that the arbitrary vector $\mathbf{x}$ in our solution does not contribute. But in many situations this condition is not or not well enough fulfilled, which can cause problems in the traditional charge fitting schemes. Although it is rare to have a matrix $\mathbf{A}$ that is strictly rank-deficient, it happens frequently that the columns of $\mathbf{A}$ are

almost linearly dependent, which is sufficient to cause numerical difficulties in the evaluation of eq. (26). This problem is aggravated by the fact that, in multiplying the original equation by $\mathbf{A}^T$, one effectively squares the condition number of the linear equations.

The physical reason for rank deficiency in the matrix **A** is that the values of the potential at the chosen points are not sufficient to determine the ''best'' charges unambiguously; i.e., there are groups of charges (corresponding to the columns of **A** that are almost linearly dependent) whose combined contribution to the potential is very small at all evaluation points. Consequently, the error measure, χ^2, is not sensitive to changes in the values of these charges. Numerical instabilities in eq. (26) can then cause very large fictitious values for these charges that lead to problems when these charges are used for other configurations or at other evaluation points. Bayly et al.[28] have imposed restraints on the absolute magnitude of the charges to eliminate such unphysically high values.

In our approach this difficulty is alleviated by calculating the generalized inverse with a numerically stable method; we have used singular value decomposition (SVD), which is the preferred method for calculating generalized inverses. In the case of strict rank deficiency, choosing the minimal length solution in eq. (24) by setting $\mathbf{x} = 0$ ensures that charges that are not well defined receive the smallest values consistent with the fit. For our application, no further measures had to be taken to ensure a reasonable fit.

There are, however, cases in which even SVD produces erratic results, namely when the condition number of **A** is too high. Therefore the condition number should always be checked in such calculations. The easiest solution in situations with high condition numbers is to set the smallest singular values to zero[31]; this corresponds physically to eliminating the charge groups whose effects almost cancel out. Alternatively, one can try to find and add evaluation points at which the potential of these groups is clearly non-zero, or another conformation of the molecule whose electrostatic potential is more sensitive to their contribution.

Acknowledgments

This work was supported by the FCAR. B. R. is a FRSQ research fellow. K. H. was a post-doctoral fellow of the Groupe de Recherche en Transport Membranaire (GRTM).

References

1. R. P. Bell, *The Proton in Chemistry*, 2nd ed., Cornell University Press, Ithaca, NY, 1973.
2. R. Stewart, *The Proton: Applications to Organic Chemistry*, Academic Press, Orlando, FL, 1985.
3. H. Azzouz and D. Borgis, *J. Chem. Phys.*, **98,** 7361 (1993).
4. D. Laria, G. Ciccotti, M. Ferrario, and R. Kapral, *Chem. Phys.*, **180,** 181 (1994).
5. D. Borgis and J. T. Hynes, *J. Chem. Phys.*, **94,** 3619 (1991).
6. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965.
7. D. Chandler, In D. Levesque, J.-P. Hansen, and J. Zinn-Justin, Eds., *Liquides, Cristallisation et Transition Vitreuse, Les Houches, Session LI*. Elsevier, New York, 1991.
8. C. Leforestier, R. H. Bisseling, C. Cerjan, M. D. Feit, R. Friesner, A. Guldberg, A. Hammerich, G. Jolicard, W. Karrlein, H.-D. Meyer, N. Lipkin, O. Roncero, and R. Kosloff, *J. Comp. Phys.*, **94,** 59 (1991).
9. W. J. Hehre, L. Radom, P. v. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbit Theory*, Wiley, New York, 1986.
10. R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.
11. R. Car and M. Parrinello, *Phys. Rev. Lett.*, **55,** 2471 (1985).
12. U. L. H.-P. Cheng and R. N. Barnett, *Chem. Phys. Lett.*, **237,** 161 (1995).
13. J. Gao, *J. Chem. Soc.*, **117,** 8600 (1995).
14. A. Warshel and R. M. Weiss, *J. Am. Chem. Soc.*, **102,** 6218 (1980).
15. A. Warshel, *Computer Modelling of Chemical Reactions in Enzymes and Solutions*, Wiley, New York, 1991.
16. J. Åqvist and A. Warshel, *J. Mol. Biol.*, **224,** 7 (1992).
17. Y.-T. Chang and W. H. Miller, *J. Phys. Chem.*, **94,** 5884 (1990).
18. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.*, **4,** 187 (1983).
19. M. J. Frisch, M. Head-Gordon, G. W. Trucks, J. B. Foresman, H. B. Schlegel, K. Raghavachari, M. Robb, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, S. Topiol, and J. A. Pople. *Gaussian 90, Rev. J.* Gaussian, Inc., Pittsburgh, PA (1990).
20. S. Scheiner and E. A. Hillenbrand, *Proc. Natl. Acad. Sci. USA*, **82,** 2741 (1985).
21. H. A. Carlson, T. B. Nguyen, M. Orozco, and W. L. Jorgensen, *J. Comput. Chem.*, **14,** 1240 (1993).
22. A. D. MacKerell, Jr., D. Bashford, M. Bellot, R. L. Dunbrack, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, D. T. Nguyen, T. Ngo, B. Prodhom, B. Roux, B. Schlenkrich, J. Smith, R. Stote, J. Straub, J. Wiorkiewicz-Kuczera, and M. Karplus, *Biophys. J.*, **61,** A143 (1992).

23. F. A. Momany, *J. Phys. Chem.*, **82,** 592 (1978).
24. S. R. Cox and D. E. Williams, *J. Comput. Chem.*, **2,** 304 (1981).
25. U. C. Singh and P. A. Kollman, *J. Comput. Chem.* **5,** 129 (1984).
26. L. E. Chirlian and M. M. Francl, *J. Comput. Chem.*, **8,** 894 (1987).
27. C. M. Breneman and K. B. Wiberg, *J. Comput. Chem.*, **11,** 361 (1990).
28. C. I. Bayly, P. Cieplak, W. D. Cornell, and P. A. Kollman, *J. Phys. Chem.*, **97,** 10269 (1993).
29. R. P. Muller and A. Warshel, *J. Phys. Chem.*, **99,** 17516 (1995).
30. A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, Wiley, New York, 1974.
31. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed., Cambridge University Press, Cambridge, 1992.